Bibliometric Analysis of Tumor Immunotherapy Studies

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

ABCDEF 1,2 Kaining Lu
BCD 3 Shan Yu
DF 4 Meng Yu
CD 1,2 Dan Sun
CD 1,2 Zhihong Huang
CD 1,2 Haotian Xing
D 1 Jianbin Bi
D 1 Zeliang Li
D 1 Zhenhua Li
D 1 Xiankui Liu
A 1,2 Chuize Kong
AG 1,2 Yuyan Zhu

1 Department of Urology, The First Hospital of China Medical University, Shenyang, Liaoning, P.R. China
2 Joint Laboratory of Artificial Intelligence and Precision Medicine of China Medical University and Northeastern University, Shenyang, Liaoning, P.R. China
3 Department of Cell Biology, Key Laboratory of Cell Biology, Ministry of Public Health, Key Laboratory of Medical Cell Biology, Ministry of Education, College of Basic Medical Science, China Medical University, Shenyang, Liaoning, P.R. China
4 Department of Reproductive Biology and Transgenic Animals, China Medical University, Shenyang, Liaoning, P.R. China

Corresponding Authors: Yuyan Zhu, e-mail: yyzhu@cmu.edu.cn, Chuize Kong, e-mail: kongchuize_cmu@sina.cn
Source of support: This work was supported in part by the National Natural Science Foundation of China (81672523, 81472404, 81272834)

Background: Cancer immunotherapy is the use of the immune system to treat cancer. After years of research, there have been a significant number of publications in this field. We analyzed the literature and performed a hotspot analysis to identify important areas of future scientific research.

Material/Methods: Articles (2945) related to cancer immunotherapy published in the past 3 years were selected as the research sample. BICOMB software was then used to retrieve the high-frequency words and construct a text/co-word matrix. Next, gCLUTO software was used to analyze the matrix by double-clustering and visual analysis, in a strategy of hotspot identification.

Results: We constructed a text and co-word matrix composed of 40 high-frequency words and 2945 articles and generated a hotspot "peak map" based on double-clustering analysis. The strategic coordinates were set by use of a co-word matrix and clustering analysis. The distribution of organs or disease and the subclass of cancer immunotherapy were analyzed.

Conclusions: In this study, we classified the hot-spots of "tumor immunotherapy" into 6 categories and 8 aspects. Calculation and analysis revealed that the field of tumor immunotherapy shows a slight trend of polarization, and the immune checkpoint inhibitor PD1 blocker shows the greatest potential for future development.

MeSH Keywords: Bibliometrics • Cancer Vaccines • Immunotherapy • Neoplasms by Site • Programmed Cell Death 1 Receptor

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/910724

📄 3298   ▦ 6   📊 6   📚 25

Background

With increased understanding of the molecular mechanism of tumor development and the rapid development of biotechnology, immunotherapy has become an important part of cancer treatment together with surgery, radiotherapy, and chemotherapy [1]. Immunotherapy involves use of the immune system to treat cancer, and important efforts have been made to use immunotherapy in the treatment of many kinds of tumors in recent years, including immunological checkpoint inhibitors [2], tumor vaccines [3], CAR-T (Chimeric Antigen Receptor T Cell Immunotherapy) [4], and adoptive immunotherapy [5].

The increased study of tumor immunotherapy has resulted in many publications of basic research, clinical trials, and curative effect evaluation. These articles are an important information resource and are helpful to understand the trend of development of the field and to predict future trends. However, to the best of our knowledge, there has been no integrated data analysis of the available literature resources in the field of cancer immunotherapy.

The growing availability of bibliometric methods and tools enables collection and analysis of appropriate literature resources to judge the development status of a discipline and predict its development prospects. Since researchers publish more articles about topics that many find important, analysis can reveal research hot-spots and trends. Co-word analysis is an important method of bibliometric analysis that can identify the trends and hot topics of a subject [6]. In a single article, if 2 terms co-occur, then these 2 terms may have a potential relationship. If 2 terms co-occur frequently in the same paper, it means that they are closely related. Using the "relationships" between terms as measured by co-occurrence, statistical methods such as cluster analysis and factor analysis can then be applied. Keywords that meet preset thresholds can be considered a research hotspot based on the field and content [7]. Clustering analysis can be used to obtain the semantic relationships of the research topic. Compared with traditional clustering methods [8], double-clustering analysis can cluster the rows and columns of a matrix simultaneously, can easily cluster global information, and can be used to analyze high-dimensional data [9,10]. Using this method, Schuemie et al. analyzed the field of medical informatics and predicted the development trend [11]. Cui et al. identified hot-spots in the field of Internet health information-seeking behaviors. [7].

In this study, we conducted a bibliometric analysis by co-word analysis and visualization on the topic of tumor immunotherapy. This analysis allowed assessment of the current research status of tumor immunotherapy and prediction of its development trend. We established the hot spot strategic coordinates of tumor immunotherapy and constructed a distribution of relevant articles based on the affected organs or disease to guide project selection, research design, and project evaluations for funding institutions.

Material and Methods

This study utilized the PubMed database, a comprehensive medical database developed and operated by the National Biotechnology Information Center in the United States. The PubMed medical literature data are mainly from the MEDLINE database, and include more than 25 million articles. To find accurate and relevant data, we used the "Mesh Major Topic" of the advanced search function of PubMed to retrieve articles related to "tumor immunotherapy" and published in the last 3 years. Articles related to transplantation immunity were excluded. The search details are as follows: ("neoplasms" [MeSH Major Topic] AND "immunotherapy" [MeSH Major Topic]) NOT "transplantation" [MeSH Terms] AND ("2014-9-11" [PDat]: "2017-9-11["PDat"]). The search resulted in 2945 related papers that were used as the experimental literature set. These data were saved in XML format.

Data extraction and analysis

Data extraction and matrix construction were performed using co-occurrence matrix generation software (BICOMB), developed by Professor Lei Cui of China Medical University [12]. To explore the research focus of tumor immunotherapy, after data extraction, we visualized the most frequent major Medical Subject Headings term (MeSH terms) (Table 1). Using biclustering, we determined the relationship between the high-frequency major MeSH terms and the source literature.

Next, a binary matrix was generated using the source literature set generated by BICOMB as a row and the high-frequency MeSH terms as a column (Table 2). We then used gCLUTO 1.0 software, a graphical cluster toolkit developed by Rasmussen, Newman, and Karypis at the University of Minnesota [13]. Using gCLUTO and the appropriate biclustering for the analysis of double clusters, visual analysis and double-clustering analysis were performed [15]. We used repeated dichotomy as the clustering method, the cosine for the similarity function, and I2 for the criterion function. To determine the best number of clusters, we repeated the biclustering several times by selecting different numbers of clusters [14]. The lowest average similarity between classes (Esim) and the highest similarity within class (ISIMM) values can be used as the optimization results (Table 3). Biclustering visualization was then done to generate a high-frequency MeSH terms peak map (Figure 1) and a biclustering heat map (Figure 2). Based on semantic analysis, the MeSH terms and contents of representative articles in clustering were analyzed to obtain the final framework of tumor immunotherapy hot-spots.

Table 1. High-frequency major MeSH terms from the included publications on tumor immunotherapy (n=2950).

No.	MeSH terms	Frequency n (%)		Cumulative percentage, %
1	Immunotherapy/methods	1134	(8.12)	8.1296
2	Neoplasms/therapy	666	(4.77)	12.9042
3	Immunotherapy	618	(4.43)	17.3346
4	Neoplasms/immunology	308	(2.20)	19.5426
5	Immunotherapy, adoptive/methods	265	(1.89)	21.4424
6	Antineoplastic agents/therapeutic use	228	(1.63)	23.0769
7	Antibodies, monoclonal/therapeutic use	224	(1.60)	24.6828
8	T-Lymphocytes/immunology	204	(1.46)	26.1452
9	Melanoma/therapy	152	(1.08)	27.2349
10	Cancer vaccines/therapeutic use	116	(0.83)	28.0665
11	Cancer vaccines/immunology	111	(0.79)	28.8623
12	Immunotherapy, adoptive	105	(0.75)	29.6150
13	Neoplasms/drug therapy	96	(0.68)	30.3032
14	Antigens, neoplasm/immunology	96	(0.68)	30.9915
15	Lung neoplasms/therapy	94	(0.67)	31.6654
16	Immunotherapy/trends	92	(0.65)	32.3249
17	Uterine cervical neoplasms/prevention & control	87	(0.62)	32.9486
18	Dendritic cells/immunology	87	(0.62)	33.5723
19	Papillomavirus infections/prevention & control	85	(0.60)	34.1817
20	Brain neoplasms/therapy	84	(0.60)	34.7839
21	Antineoplastic combined chemotherapy protocols/therapeutic use	80	(0.57)	35.3574
22	Receptors, antigen, T-Cell/immunology	79	(0.56)	35.9237
23	T-Lymphocytes/transplantation	73	(0.52)	36.4471
24	Molecular targeted therapy/methods	70	(0.50)	36.9489
25	Killer cells, natural/immunology	67	(0.48)	37.4292
26	Immunotherapy/adverse effects	60	(0.43)	37.8593
27	Melanoma/drug therapy	60	(0.43)	38.2895
28	Skin neoplasms/therapy	59	(0.42)	38.7125
29	T-lymphocytes, cytotoxic/immunology	58	(0.41)	39.1283
30	Glioblastoma/therapy	57	(0.40)	39.5369
31	Programmed cell death 1 receptor/antagonists & inhibitors	54	(0.38)	39.9240
32	Carcinoma, non-small-cell lung/therapy	52	(0.37)	40.2968
33	Melanoma/immunology	52	(0.37)	40.6696
34	Vaccination	52	(0.37)	41.0424
35	Radioimmunotherapy/methods	52	(0.37)	41.4152
36	CD8-positive T-lymphocytes/immunology	51	(0.36)	41.7808
37	Molecular targeted therapy	50	(0.35)	42.1392
38	Liver neoplasms/therapy	44	(0.31)	42.4547
39	Antineoplastic agents/administration & dosage	44	(0.31)	42.7701
40	Carcinoma, hepatocellular/therapy	39	(0.27)	43.0497

Table 2. High-frequency major MeSH terms-source articles matrix (localized).

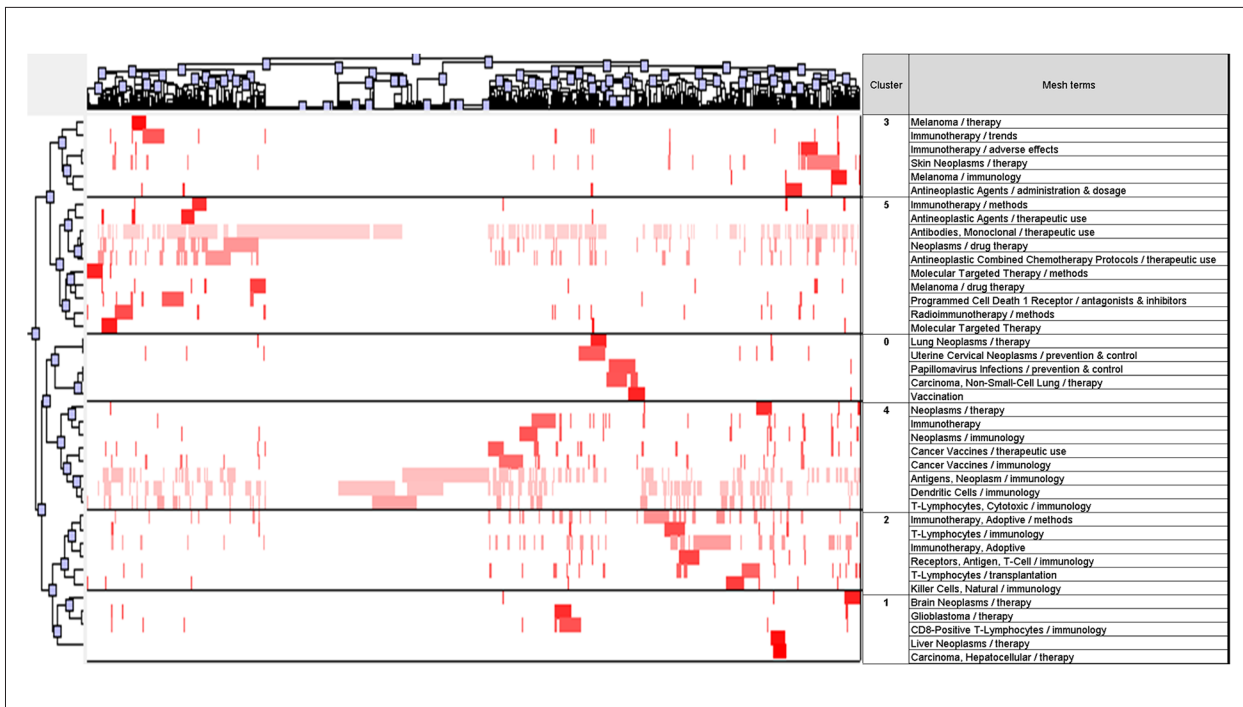| No.	Major MeSH terms	PubMed Unique Identifiers of source article				
		24051308	24141774	24240689	...	25440606
1	Immunotherapy/methods	0	0	1	...	0
2	Neoplasms/therapy	1	1	0	...	0
3	Immunotherapy	1	1	0	...	0
4	Neoplasms/immunology	0	0	0	...	0
...	...	...	...	...	...	...
39	Antineoplastic agents/administration & dosage	0	0	0	...	0
40	Carcinoma, hepatocellular/therapy	0	0	0	...	0

Table 3. Descriptive and discriminating features and representative articles.

| Descriptive and Descriminating features			
Cluster 0	**Size 5**	**ISim: 0.365**	**Esim: 0.013**
Descriptive	26826045	25300588	26845632	27488178
Descriminating	26868944	27488178	26826045	26291437
Cluster 1	**Size 5**	**ISim: 0.329**	**Esim: 0.021**
Descriptive	27188748	26059190	24691976	25929570
Descriminating	27188748	24691976	26059190	27410037
Cluster 2	**Size 6**	**ISim: 0.279**	**Esim: 0.029**
Descriptive	26998495	26384559	28347250	26332003
Descriminating	26384559	26998495	26332003	28347250
Cluster 3	**Size 6**	**ISim: 0.240**	**Esim: 0.027**
Descriptive	27259558	27302423	25640488	26837440
Descriminating	27259558	27123564	27769635	26181188
Cluster 4	**Size 8**	**ISim: 0.229**	**Esim: 0.038**
Descriptive	27706668	25196947	25904595	26923002
Descriminating	27706668	25196947	25904595	26923002
Cluster 5	**Size 10**	**ISim: 0.178**	**Esim: 0.029**
Descriptive	28387388	27710977	26578726	27438833
Descriminating	28387388	27710977	26578726	26250412

Strategic coordinates

Co-word analysis can be used as a tool to understand and describe the relationship between scientific topics. Co-word analysis can help to distinguish the local environment and the global environment of each research topic [15]. By using Excel, the co-word matrix composed of high-frequency words was used to calculate the intra-class link averages and the inter-class link average (Table 4), allowing calculation of centrality and density, respectively (Table 5). Using two-dimensional coordinates with centrality and density as parameters, a graph was constructed to describe the internal integrality of certain topics and the effects of their interactions. In a strategic diagram, the intensity of the interaction of topics is expressed with the

Figure 1. Visualized matrix of biclustering of highly frequent major MeSH terms and PubMed Unique Identifiers (PMIDs) of articles on cancer immunotherapy.
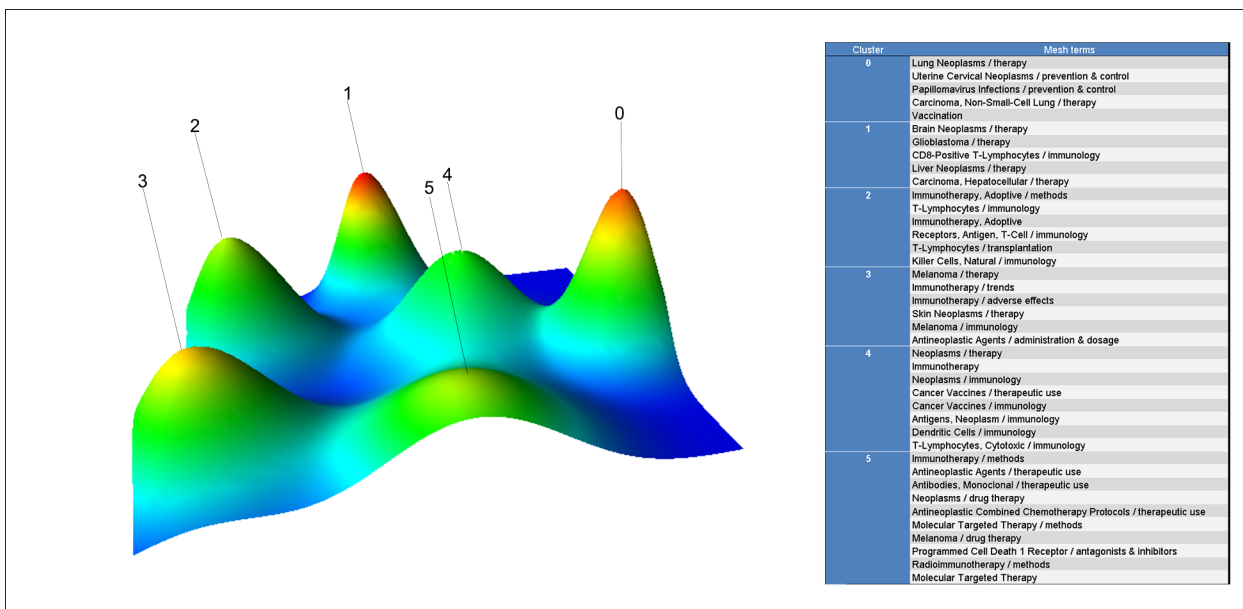
Figure 2. Mountain visualization of biclustering of highly frequent major MeSH terms and articles on tumor immunotherapy.

x-axis as the centrality. The greater the number and intensity of the links between one subject area and other disciplines, the more central the subject area is to the overall research. The centrality of a category is calculated by the strength of the links between the main items of the category and other categories. The Y-axis represents the density, indicating the strength of the internal integrality within a given category,

and the level of each category can maintain and develop itself [16]. Law used the method of co-word and strategic coordinate analysis to predict the future trend for the domain of environmental acidification research [6] Zhao et al. used this method to analyze the theme trends and knowledge structure of studies on choroidal neovascularization [17].

Table 4. A co-word matrix of high-frequency major MeSH terms (localized).

No.	Mesh terms	Immunotherapy/methods	Neoplasms/therapy	...	Carcinoma, hepatocellular
1	Immunotherapy/methods	1134	308	...	18
2	Neoplasms/therapy	308	666	...	0
...	...	...	...	...	...
40	Carcinoma, hepatocellular	18	0	...	39

Table 5. The centrality and density of the 6 clusters.

| Cluster	Intra-class link averages	Density-Y	Inter-class link average	Centralit-X
0	34	−4.66	1.68	−3.01
1	22.25	−16.41	2.52	−2.17
2	44.03	5.37	4.92	0.228
3	22.7	−15.96	3.57	−1.12
4	68.5	29.91	8.55	3.86
5	40.42	1.76	6.9	2.21
Average	38.65		4.69	

Results

For articles published from 11 September 2014 to 11 September 2017, 2950 major topic terms were extracted. In the high-frequency term table (Table 1), the frequency number before the 39th word is greater than its ordinal number, and the frequency number of the high-frequency term after the 40th word is less than its ordinal number. Thus, the terms ranked higher than 40 can be defined as very frequent, with a cumulative number of 5878. Forty high-frequency terms were then extracted from the included publications, with a cumulative percentage of 42.05% (Table 1). According to the co-occurrence of these high-frequency terms in the same article, we established a co-occurrence matrix of these terms with the major topic terms as rows and the source articles as columns (Table 2). The matrix includes the major topic terms in the source articles (Table 2). A "1" in a cell indicates that the term is present in the article, and "0" means that it is not present. Then, a matrix of co-words was established (Table 4). In that matrix, the numbers indicate the times of co-occurrence between 2 terms (Table 4).

To describe the hot topic distribution in the field of tumor immunotherapy, we used the Carrot system to search for articles, an online visualization system based on PubMed. We searched for "cancer immunotherapy" and "tumor immunotherapy". In the PubMed database, the larger the weight of the topic, the larger the area will be, and the more central the module (Figure 3).

We used different parameters to double-cluster the matrix and then we performed visualization by gCLUTO to generate a peak map and heat map using the matrix of high-frequency major topic terms and source paper (Table 2). The purpose of the visualization of the heat map is to understand the content of the high-dimensional data set more intuitively. In Figure 2, each category is numbered from 0 to 5. Peak, volume, height, and color are all used to depict information about the associated cluster under the preset condition. The position of a peak is its vertical distance from the top to the plane surface. The relative similarity of clusters is shown as the distance between 2 peaks on the plane. The height of each peak is proportional to the internal similarity of the cluster. The volume of the peak is proportional to the quantity of major MeSH terms for each cluster. Finally, the color of the peak domain indicates the standard deviation within the cluster's objects, where red means low deviation and blue indicates high deviation. Figure 1 shows the heat map of double-clustering visualization, where rows represent the high-frequency major MeSH terms, with the columns of corresponding terms located on the right. The bottom of the matrix shows the PubMed unique identifier of each source article. The depth of color for each grid indicates the relative frequency of the major MeSH terms in the article. A larger value will show a deeper red grid, white represents a value closer to zero, and negative values are green. The double-clustering matrix visualization showed that 40 highly frequent major MeSH terms were clustered in 6 peaks. In Figure 1, the hierarchical tree on the left edge depicts the relationship between these high-frequency MeSH terms,
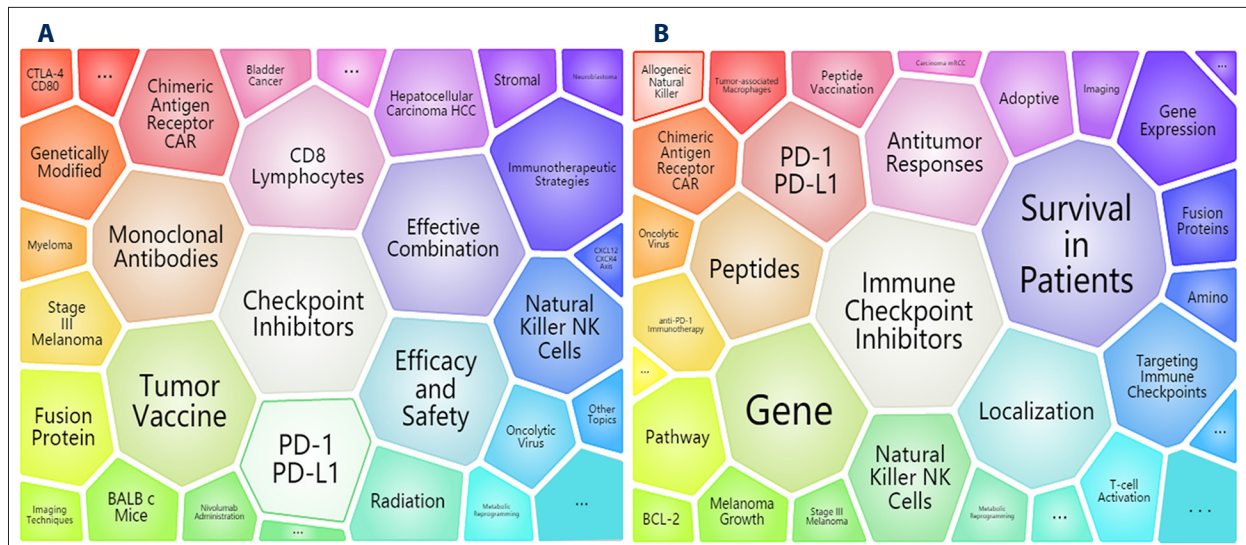
Figure 3. A survey of major topics in tumor immunotherapy and cancer immunotherapy is shown. The visualizations were done using the Carrot system based on the top-ranking results of search. **(A)** Search results of tumor immunotherapy. **(B)** Search results of cancer immunotherapy.

and the top-hierarchical tree shows the relationship between articles. The tree shows each high-frequency MeSH terms in the source literatures for each category. This presents the semantic relationship between the subject terms and articles, allowing identification and summarization of the themes of each category (Table 3).

In addition, according to the following criteria discussed by the study group of 2 researchers, some clusters can be subdivided into complete or smaller topics: 1) The semantic relationship between the major MeSH terms in a large cluster, 2) Category in which these words were clustered, and 3) Representative articles (Table 3) in each category. These allow smaller themes to be grouped into a single hotspot. We found 6 clusters of hot-spots of ideas in cancer immunotherapy, including 8 hot topics, as follows:

1. Immunotherapy for lung cancer (cluster 0).
2: Vaccine for cervical cancer (cluster 0).
3: Immunotherapy for hepatocellular carcinoma (cluster 1).
4: Immunotherapy for brain tumor (glioma) (cluster 1).
5: Adoptive immunotherapy based on natural killer T cells and their proliferation (cluster 2).
6: Immunotherapy for melanoma and skin neoplasms, including drug dosage and adverse effects of drugs (cluster 3).
7: Dendritic cell vaccine and other cancer treatment vaccines associated with T lymphocyte cytotoxicity (cluster 4).
8: Immune checkpoint inhibitors such as the PD1 blocker (cluster 5).

The horizontal axis of the strategic coordinate indicates the centrality, the vertical axis represents the density, and the first quadrant is the upper-right corner, then moving clockwise, the
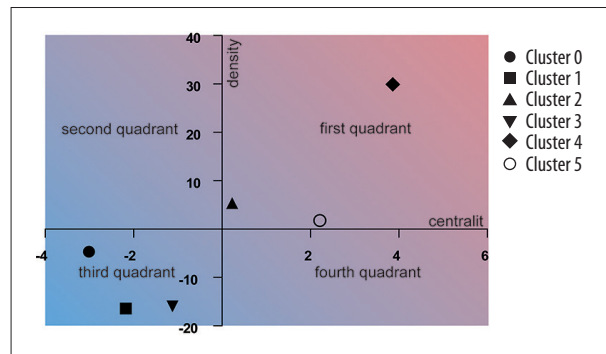
Figure 4. Strategic diagram of clusters.

second quadrant, the third quadrant, and the fourth quadrant. The darker the red, the higher the value of centrality and density, and the darker the blue, the lower the value of centrality and density. As we can see from Figure 4, clusters 2, 4, and 5 are in the first quadrant, representing the corresponding category in the central and core field. Clusters 0, 1, and 3 are in the third quadrant, indicating that their corresponding categories are in relatively peripheral cold fields (Figure 4).

After the above macroscopic analysis, we further judged the distribution and prevalent trends of tumor immunotherapy at the level of affected organ or disease. We selected the first 1000 articles from the source articles and classified the articles by the diseases or organs studied, resulting in 594 articles that were related to organs or specific diseases. Using to the map of organ distribution, one can assess the general situation of tumor immunotherapy and its relatedness to the study of each organ or disease in the last 3 years (Table 6, Figure 6).
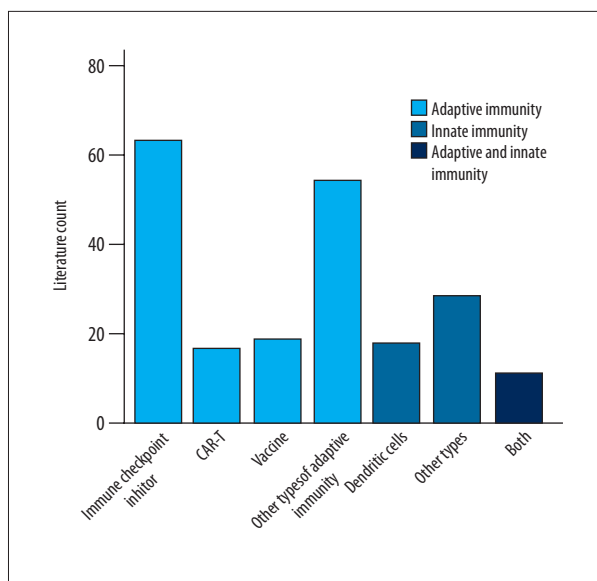
Figure 5. Weight and distribution of each subcategory in tumor immunotherapy.

Similarly, by using a single article as a unit, we analyzed the current status and hot-spots of tumor immunotherapy subgroups, and our team extracted and studied the first 300 abstracts of source articles to further construct subgroups in this field. At the same time, we evaluated whether the type of immunity was innate, adaptive, or both. The research team set the following criteria for classification: if the article focused on adaptive immunity, such as PD1/PDL1/CAR-T, we classified it as an adaptive immune category; if the article focused on the innate immune molecules, cells, pathways such as IFN, dendritic cells, macrophages, and related processes, we classified it as innate immunity. From the sample set containing 300 articles, 87 articles without mention of innate or adaptive immunity were excluded. Of the remaining 213 articles, there were 155 articles related to adaptive immunization, including immune checkpoint inhibitors (64/155), CAR-T (17/155), cancer vaccines, (19/155), and other types (55/155) such as immunized T cells or CD-27. There were 47 articles related to innate immunity, including dendritic cells (18/47) and other molecules, cells, and pathways, and 11 articles on both innate and adaptive immunity. We then predicted the weight and distribution of each subcategory in tumor immunotherapy (Figure 5).
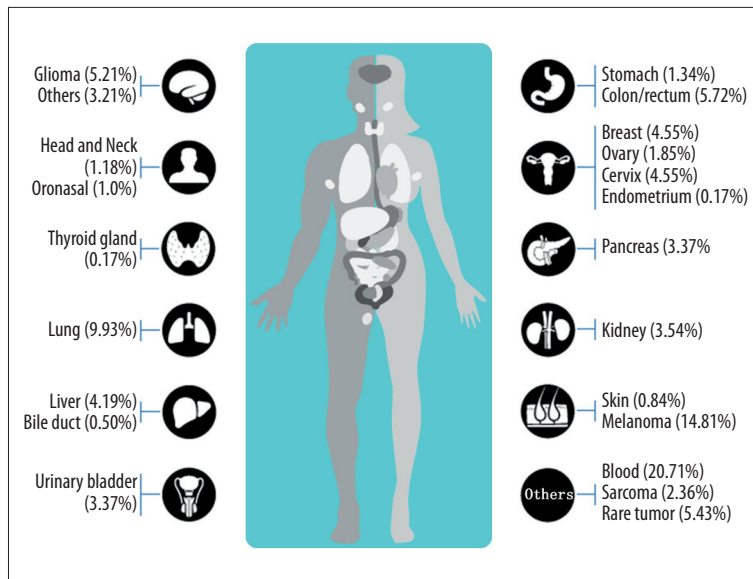
Discussion

Tumor immunotherapy is a promising field in cancer treatment [18]. There have been many research papers and reviews of basic research findings and clinical research. However, these studies often focus on topics such as specific genes, pathways, or other aspects and there has been no macroscopic analysis or scientific prediction of tumor immunotherapy

using bibliometrics. In this study, we used the method of co-word research and double-clustering visualization analysis to obtain 6 categories of 8 aspects of tumor immunotherapy as hot-spots. According to the co-words matrix, we constructed hotspot strategic coordinates, generated an organ or disease distribution map to understand the weight of each organ or disease in the tumor immunotherapy field, and generated statistics for the tumor immune subcategory. These results suggest that the study of cancer vaccines is an absolute core topic with much interaction with the other categories and good integrality; the ability of adoptive immunotherapy to maintain and develop itself is limited, and the immune checkpoint inhibitor PD1 blocker is the most promising (Table 5, Figure 4).

For tumor immunotherapy, we constructed hot strategic coordinates based on co-word analysis. In the strategic coordinate plot, clusters 2, 4, and 5 are in the first quadrant, indicating relatively hot spots, and clusters 0, 1, and 3 are in the third quadrant, which means relatively cold spots. The 3 topics in the core area are adoptive immunotherapy, cancer vaccines, and PD1-targeted therapy, meaning that these topics are relatively central in the current research environment. These 3 topics are in the first quadrant, but their positions in the coordinate system are different, so it is very important to analyze the relationships among them. Cluster 4 (cancer vaccines) is at the core because of the high centrality and density of the cluster in its category. The high centrality value indicates that it is closely related to the other hot-spots, and the high density indicates that the topics are inclined to be self-integrated and the research tends to be mature. Cluster 2 (adoptive immunotherapy) and cluster 5 (PD1 blocker) are located very closely to the axis. Although the positions of these clusters are in the core field, they are located at the edge, suggesting that cluster 2 (accepted immunotherapy) and cluster 5 (PD1 blocker) are not as hot as cluster 4 (cancer vaccine). Cluster 5 (PD1 blocker) has a high degree of centrality and a low density, which means that it is closely related to other hot-spots, but its internal integrality is not tight enough and the research is not fully mature. Cluster 2 (adoptive immunotherapy) has a high density and a relatively low degree of centrality, suggesting that the topics are well-integrated and the research is mature, but not closely related to the other hotspot categories. Because of the high centrality value, the potential of cluster 5 (PD1 blocker) is higher than that of cluster 2 (advanced immunotherapy) and is the most valuable of the clusters. Clusters 0, 1, and 3 are in the third quadrant, possibly because these clusters have independent topics that do not have significant connections with each other. The second quadrant represents an isolated field, which indicates that the domain research has a large independence and temporary scale but lacks connection with the overall field. The fourth quadrant represents potential areas where the field research is immature but closely related to the overall field, and so has greater

Figure 6. Affected organ or disease distribution of tumor immunotherapy.

Glioma (5.21%)
Others (3.21%)

Head and Neck (1.18%)
Oronasal (1.0%)

Thyroid gland (0.17%)

Lung (9.93%)

Liver (4.19%)
Bile duct (0.50%)

Urinary bladder (3.37%)

Stomach (1.34%)
Colon/rectum (5.72%)

Breast (4.55%)
Ovary (1.85%)
Cervix (4.55%)
Endometrium (0.17%)

Pancreas (3.37%

Kidney (3.54%)

Skin (0.84%)
Melanoma (14.81%)

Blood (20.71%)
Sarcoma (2.36%)
Rare tumor (5.43%)

Table 6. Literature count of affected organ or disease.

Organ/disease & literature count		
Brain tumor (19/594)	Lung cancer (59/594)	Cervix cancer(27/594)
Glioma (31/594)	Liver cancer (25/594)	Skin cancer (5/594)
Cancer of head & neck (7/594)	Thyroid gland (1/594)	Myeloma (9/594)
Pancreatic cancer (20/594)	Bladder cancer (20/594)	Lymphoma (35/594)
Colon & rectum (34/594)	Prostate cancer (12/594)	Leukaemia (50/594)
Stomch (8/594)	Renal carcinoma (21/594)	Other blood disease(27/594)
Mallanoma (88/594)	Endometrium (1/594)	Oronasa cancer (6/594)
Cholangio carcinoma (3/594)	Ovary cancer(11/594)	Oral tumor (2/594)
Sarcoma (14/594)	Breast cancer(27/594)	Rare tumor (32/594)

potential importance. No category appears in the second and fourth quadrants, which may indicate that the hot-spots in the field of tumor immunotherapy show a slight polarization trend, which may mean this distribution of the discipline will not remain unchanged.

Hot-spots and cold-spots tend to change dynamically. If researchers pursue work in absolute hot-spots, there may be an increase in the volume of hot-spots, leading to "overheating", and this evolution may cause the hotspot to undergo a conversion. For example, with the increase of cluster 5 (PD1 blocker) research, an increase in the degree of association of adoptive immunotherapy (cluster 2) with other fields may result in the replacement of cluster 4 (cancer vaccine) by these 2 clusters in the core mature position. The 3 clusters of the third quadrant may move to other quadrants as the volume increases and the association with other clusters increases.

The organ or disease profile (Figure 6) shows that the 2 most important organs or diseases for tumor immunotherapy are the blood system and melanoma. This distribution may be due to the higher efficacy of CAR-T and PD1 on hematologic tumors and melanoma, respectively [19–22]. There were few counts for several tumors with poor prognosis such as pancreatic cancer and gastric cancer, and this may due to their limited response to immunotherapy [23]. With in-depth study of the mechanism of tumor immunotherapy, there is opportunity for study of malignant tumors with poor prognosis.

Analysis of the extracted literature set showed that recent cancer immunotherapy research is dominated by adaptive immunotherapy (Figure 5). In the study of adaptive immunity, subcategory analysis of tumor immunotherapy indicated that the most popular topics were immune checkpoint inhibitors, CAR-T, and cancer vaccines, and the most important topic in innate

immunity research was dendritic cells, likely due to the ability of these cells to induce primary immune responses and connect innate and adaptive immunity systems [24]. Innate immunity is the first line of defense against cancer and also assists adaptive immunity, which has great exploration space [25].

Although co-word clustering analysis of high-frequency words is a new method of analysis, there may be a certain degree of bias due to word selection among researchers when writing. Additionally, the quality of articles in the PubMed database is not uniform, and although 2945 articles were examined, this dataset may be incomplete due to limitations of intelligence in the retrieval system.

Conclusions

For tumor immunotherapy, we summarized 6 categories and 8 hot topics. The current state of research in this field is polarized; cancer vaccines are at the absolute core with the most mature research, and the ability of adoptive immunotherapy to sustain and develop itself is limited. The immune checkpoint inhibitor (PD1 blocker) is very promising. Citation frequency analysis is also an important method for bibliometric research, and citation analysis of tumor immunotherapy articles should be done in the future.

Acknowledgements

References:

1. Mellman I, Coukos G, Dranoff G: Cancer immunotherapy comes of age. Nature, 2011; 480(7378): 480–89

2. Pardoll DM: The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer, 2012; 12(4): 252–64

3. Rosenberg SA, Yang JC, Restifo NP: Cancer immunotherapy: Moving beyond current vaccines. Nat Med, 2004; 10(9): 909–15

4. Daher M, Rezvani K: Next generation natural killer cells for cancer immunotherapy: The promise of genetic engineering. Curr Opin Immunol, 2018; 51: 146–53

5. Rosenberg SA, Restifo NP, Yang JC et al: Adoptive cell transfer: A clinical path to effective cancer immunotherapy. Nat Rev Cancer, 2008; 8(4): 299–308

6. Law J, Bauin S, Courtial J-P, Whittaker J: Policy and the mapping of scientific change: A co-word analysis of research into environmental acidification. Scientometrics, 1988; 14(3–4): 251–64

7. Li F, Li M, Guan P et al: Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: A quantitative and co-word biclustering analysis. J Med Internet Res, 2015; 17(3): e81

8. Zhang Z, Murtagh F, Van Poucke S et al: Hierarchical cluster analysis in clinical research with heterogeneous study population: Highlighting its visualization with R. Ann Transl Med, 2017; 5(4): 75

9. Cheng Y, Church GM: Biclustering of expression data. Eighth International Conference on Intelligent Systems for Molecular Biology AAAI Press, 2000; 8: 93–103

10. Hartigan JA: Direct clustering of a data matrix. Publications of the American Statistical Association, 1978; 67(337): 123–29

11. Schuemie MJ, Talmon JL, Moorman PW, Kors JA et al: Mapping the domain of medical informatics. Methods Inf Med, 2009; 48(1): 76–83

12. Cui L, Liu W, Yan L et al: Development of a text mining system based on the co-occurrence of bibliographic items in literature. New Technology of Library and Information Service, 2008; (8): 70–75

13. Karypis Lab. Webcite gCLUTO-Graphical Clustering Toolkit: *http://glaros. dtc.umn.edu/gkhome/cluto/gcluto/download* [accessed 2018-04-17]

14. Wang L: [Application of biclustering on classification of scientific researchers based on literature profiling Shenyang.] China Medical University (PRC), 2012; 1(1) [in Chinese]

15. Bauin S, Michelet B, Schweighoffer MG, Vermeulin P: Using bibliometrics in strategic analysis: "understanding chemical reactions" at the CNRS. Scientometrics, 1991; 22(1): 113–37

16. Callon M, Courtial JP, Laville F: Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. Scientometrics, 1991; 22(1): 155–205

17. Zhao F, Shi B, Liu R et al: Theme trends and knowledge structure on choroidal neovascularization: A quantitative and co-word analysis. BMC Ophthalmol, 2018; 18(1): 86

18. Helmy KY, Patel SA, Nahas GR, Rameshwar P: Cancer immunotherapy: Accomplishments to date and future promise. Ther Deliv, 2013; 4(10): 1307–20

19. Hamid O, Robert C, Daud A et al: Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. N Engl J Med, 2013; 369(2): 134–44

20. Patnaik A, Kang SP, Rasco D et al: Phase I study of pembrolizumab (MK-3475; anti-PD-1 monoclonal antibody) in patients with advanced solid tumors. Clin Cancer Res, 2015; 21(19): 4286–93

21. Porter DL, Levine BL, Kalos M et al: Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. N Engl J Med, 2011; 365(8): 725–33

22. Grupp SA, Kalos M, Barrett D et al: Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. N Engl J Med, 2013; 368(16): 1509–18

23. Xu JW, Wang L, Cheng YG et al: Immunotherapy for pancreatic cancer: A long and hopeful journey. Cancer Lett, 2018; 425: 143–51

24. Banchereau J, Briere F, Caux C et al: Immunobiology of dendritic cells. Annu Rev Immunol, 2000; 18: 767–811

25. Vesely MD, Kershaw MH, Schreiber RD et al: Natural innate and adaptive immunity to cancer. Annu Rev Immunol, 2011; 29(1): 235–71